

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/506, 670A  
Source: IFWO  
Date Processed by STIC: 09/12/2006

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:14

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

```

3 <110> APPLICANT: LIPKA, VOLKER
4     SCHEEL, DIERK
5     SCHULZE-LEFERT, PAUL
6     ROSAHL, SABINE
7     LANDLAG, JORN
9 <120> TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A BETA-GLUCOSIDASE AND USES THEREOF
11 <130> FILE REFERENCE: G1309 US S3
13 <140> CURRENT APPLICATION NUMBER: 10/506,670A
14 <141> CURRENT FILING DATE: 2004-09-03
16 <150> PRIOR APPLICATION NUMBER: PCT/EP03/02315
17 <151> PRIOR FILING DATE: 2003-03-06
19 <150> PRIOR APPLICATION NUMBER: EP 02004400.4
20 <151> PRIOR FILING DATE: 2002-03-06
22 <160> NUMBER OF SEQ ID NOS: 9
24 <170> SOFTWARE: PatentIn Ver. 3.3
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1683
28 <212> TYPE: DNA
29 <213> ORGANISM: Arabidopsis thaliana
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (1)..(1680)
35 <400> SEQUENCE: 1
36 atg gca cat ctt caa aga aca ttt cct act gag atg tca aaa ggt aga      48
37 Met Ala His Leu Gln Arg Thr Phe Pro Thr Glu Met Ser Lys Gly Arg
38 1          5          10          15
40 gct agc ttt cca aaa ggc ttt ctc ttt gga act gct tca tct tct tat      96
41 Ala Ser Phe Pro Lys Gly Phe Leu Phe Gly Thr Ala Ser Ser Ser Tyr
42          20          25          30
44 cag tac gaa gga gca gtg aat gaa ggt gcg aga gga caa agc gtg tgg      144
45 Gln Tyr Glu Gly Ala Val Asn Glu Gly Ala Arg Gly Gln Ser Val Trp
46          35          40          45
48 gat cat ttc tcc aac agg ttt cct cac aga atc agt gat tct agc gac      192
49 Asp His Phe Ser Asn Arg Phe Pro His Arg Ile Ser Asp Ser Ser Asp
50 50          55          60
52 gga aac gtt gcc gtt gat ttc tac cat cgt tac aag gaa gat att aag      240
53 Gly Asn Val Ala Val Asp Phe Tyr His Arg Tyr Lys Glu Asp Ile Lys
54 65          70          75          80
56 aga atg aag gat ata aac atg gat tcg ttt cgg ctt tcc att gct tgg      288
57 Arg Met Lys Asp Ile Asn Met Asp Ser Phe Arg Leu Ser Ile Ala Trp
58          85          90          95
60 cca cga gtt cta cct tat ggc aaa agg gat aga gga gtt agt gaa gaa      336
61 Pro Arg Val Leu Pro Tyr Gly Lys Arg Asp Arg Gly Val Ser Glu Glu

```

## RAW SEQUENCE LISTING

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:14

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

```

62          100          105          110
64 gga att aag ttt tac aat gat gtt att gat gaa ctc tta gcc aat gaa 384
65 Gly Ile Lys Phe Tyr Asn Asp Val Ile Asp Glu Leu Leu Ala Asn Glu
66          115          120          125
68 atc act cct ctt gtt act atc ttt cat tgg gac ata cca cag gat ctt 432
69 Ile Thr Pro Leu Val Thr Ile Phe His Trp Asp Ile Pro Gln Asp Leu
70          130          135          140
72 gaa gat gaa tac ggc ggt ttt cta agc gag cag att ata gat gac ttc 480
73 Glu Asp Glu Tyr Gly Gly Phe Leu Ser Glu Gln Ile Ile Asp Asp Phe
74 145          150          155          160
76 aga gac tat gcg agt ctc tgc ttc gag aga ttt ggg gac aga gtg agt 528
77 Arg Asp Tyr Ala Ser Leu Cys Phe Glu Arg Phe Gly Asp Arg Val Ser
78          165          170          175
80 ctg tgg tgc aca atg aat gag ccg tgg gtc tac agt gtc gcg ggc tat 576
81 Leu Trp Cys Thr Met Asn Glu Pro Trp Val Tyr Ser Val Ala Gly Tyr
82          180          185          190
84 gac aca gga agg aaa gcg cca gga cgg tgc tcc aag tat gtt aat ggc 624
85 Asp Thr Gly Arg Lys Ala Pro Gly Arg Cys Ser Lys Tyr Val Asn Gly
86          195          200          205
88 gct agt gtt gct gga atg tcg gga tac gag gca tat att gtg agc cat 672
89 Ala Ser Val Ala Gly Met Ser Gly Tyr Glu Ala Tyr Ile Val Ser His
90          210          215          220
92 aac atg ctt cta gcg cac gca gaa gca gtg gaa gtg ttt aga aaa tgt 720
93 Asn Met Leu Leu Ala His Ala Glu Ala Val Glu Val Phe Arg Lys Cys
94 225          230          235          240
96 gac cat att aaa aac gga caa att ggg att gcg cat aat cca ctt tgg 768
97 Asp His Ile Lys Asn Gly Gln Ile Gly Ile Ala His Asn Pro Leu Trp
98          245          250          255
100 tac gag cca tat gat ccg agt gat cca gat gat gtc gaa gga tgt aat 816
101 Tyr Glu Pro Tyr Asp Pro Ser Asp Pro Asp Asp Val Glu Gly Cys Asn
102          260          265          270
104 cga gct atg gac ttc atg ctt ggt tgg cat cag cat ccg act gct tgt 864
105 Arg Ala Met Asp Phe Met Leu Gly Trp His Gln His Pro Thr Ala Cys
106          275          280          285
108 gga gac tat cca gaa acg atg aag aaa tca gtt gga gat aga tta ccg 912
109 Gly Asp Tyr Pro Glu Thr Met Lys Lys Ser Val Gly Asp Arg Leu Pro
110          290          295          300
112 agt ttt aca cca gaa caa tct aag aaa ctt ata ggc tct tgc gat tac 960
113 Ser Phe Thr Pro Glu Gln Ser Lys Lys Leu Ile Gly Ser Cys Asp Tyr
114 305          310          315          320
116 gtt ggt ata aac tac tat agc tcg ctt ttc gtg aag agt atc aaa cac 1008
117 Val Gly Ile Asn Tyr Tyr Ser Ser Leu Phe Val Lys Ser Ile Lys His
118          325          330          335
120 gtg gat cct acg caa cct act tgg aga act gac caa ggc gtt gat tgg 1056
121 Val Asp Pro Thr Gln Pro Thr Trp Arg Thr Asp Gln Gly Val Asp Trp
122          340          345          350
124 atg aaa acc aac ata gat ggg aaa caa ata gca aaa caa gga gga tca 1104
125 Met Lys Thr Asn Ile Asp Gly Lys Gln Ile Ala Lys Gln Gly Gly Ser
126          355          360          365

```

## RAW SEQUENCE LISTING

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:14

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

```

128 gag tgg agt ttc aca tat cca aca gga ctc aga aac att ttg aag tat      1152
129 Glu Trp Ser Phe Thr Tyr Pro Thr Gly Leu Arg Asn Ile Leu Lys Tyr
130      370      375      380
132 gtg aaa aaa act tat ggc aat cct ccc att ctc ata act gaa aac ggg      1200
133 Val Lys Lys Thr Tyr Gly Asn Pro Pro Ile Leu Ile Thr Glu Asn Gly
134 385      390      395      400
136 tat ggt gaa gta gcg gaa cag agt cag agt ctt tat atg tac aat cct      1248
137 Tyr Gly Glu Val Ala Glu Gln Ser Gln Ser Leu Tyr Met Tyr Asn Pro
138      405      410      415
140 tca atc gac aca gag aga ttg gag tac att gaa gga cat atc cac gct      1296
141 Ser Ile Asp Thr Glu Arg Leu Glu Tyr Ile Glu Gly His Ile His Ala
142      420      425      430
144 att cat caa gcc atc cat gaa gat gga gta aga gtg gaa ggt tat tac      1344
145 Ile His Gln Ala Ile His Glu Asp Gly Val Arg Val Glu Gly Tyr Tyr
146      435      440      445
148 gta tgg tca ttg cta gat aac ttc gag tgg aac agt gga tat ggt gtg      1392
149 Val Trp Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Tyr Gly Val
150      450      455      460
152 aga tat ggt tta tat tac att gat tac aaa gat ggg ctt aga cga tac      1440
153 Arg Tyr Gly Leu Tyr Tyr Ile Asp Tyr Lys Asp Gly Leu Arg Arg Tyr
154 465      470      475      480
156 ccg aaa atg tcg gcg tta tgg ttg aaa gag ttc ttg agg ttt gat caa      1488
157 Pro Lys Met Ser Ala Leu Trp Leu Lys Glu Phe Leu Arg Phe Asp Gln
158      485      490      495
160 gaa gac gat tct tcg acg tct aag aaa gaa gag aag aaa gag agc tat      1536
161 Glu Asp Asp Ser Ser Thr Ser Lys Lys Glu Glu Lys Lys Glu Ser Tyr
162      500      505      510
164 gga aaa cag tta ttg cat tct gtt cag gac agt caa ttt gtt cat tcg      1584
165 Gly Lys Gln Leu Leu His Ser Val Gln Asp Ser Gln Phe Val His Ser
166      515      520      525
168 att aaa gac agt ggt gcg tta cct gcg gtt ttg ggg agc ttg ttc gtt      1632
169 Ile Lys Asp Ser Gly Ala Leu Pro Ala Val Leu Gly Ser Leu Phe Val
170      530      535      540
172 gtg tct gca act gtt ggt act tct ctg ttc ttc aaa gga gct aat aat      1680
173 Val Ser Ala Thr Val Gly Thr Ser Leu Phe Phe Lys Gly Ala Asn Asn
174 545      550      555      560
176 tga      1683
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 560
181 <212> TYPE: PRT
182 <213> ORGANISM: Arabidopsis thaliana
184 <400> SEQUENCE: 2
185 Met Ala His Leu Gln Arg Thr Phe Pro Thr Glu Met Ser Lys Gly Arg
186 1      5      10      15
188 Ala Ser Phe Pro Lys Gly Phe Leu Phe Gly Thr Ala Ser Ser Ser Tyr
189      20      25      30
191 Gln Tyr Glu Gly Ala Val Asn Glu Gly Ala Arg Gly Gln Ser Val Trp
192      35      40      45
194 Asp His Phe Ser Asn Arg Phe Pro His Arg Ile Ser Asp Ser Ser Asp

```

## RAW SEQUENCE LISTING

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:14

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

```

195      50      55      60
197 Gly Asn Val Ala Val Asp Phe Tyr His Arg Tyr Lys Glu Asp Ile Lys
198 65      70      75      80
200 Arg Met Lys Asp Ile Asn Met Asp Ser Phe Arg Leu Ser Ile Ala Trp
201      85      90      95
203 Pro Arg Val Leu Pro Tyr Gly Lys Arg Asp Arg Gly Val Ser Glu Glu
204      100      105      110
206 Gly Ile Lys Phe Tyr Asn Asp Val Ile Asp Glu Leu Leu Ala Asn Glu
207      115      120      125
209 Ile Thr Pro Leu Val Thr Ile Phe His Trp Asp Ile Pro Gln Asp Leu
210      130      135      140
212 Glu Asp Glu Tyr Gly Gly Phe Leu Ser Glu Gln Ile Ile Asp Asp Phe
213 145      150      155      160
215 Arg Asp Tyr Ala Ser Leu Cys Phe Glu Arg Phe Gly Asp Arg Val Ser
216      165      170      175
218 Leu Trp Cys Thr Met Asn Glu Pro Trp Val Tyr Ser Val Ala Gly Tyr
219      180      185      190
221 Asp Thr Gly Arg Lys Ala Pro Gly Arg Cys Ser Lys Tyr Val Asn Gly
222      195      200      205
224 Ala Ser Val Ala Gly Met Ser Gly Tyr Glu Ala Tyr Ile Val Ser His
225      210      215      220
227 Asn Met Leu Leu Ala His Ala Glu Ala Val Glu Val Phe Arg Lys Cys
228 225      230      235      240
230 Asp His Ile Lys Asn Gly Gln Ile Gly Ile Ala His Asn Pro Leu Trp
231      245      250      255
233 Tyr Glu Pro Tyr Asp Pro Ser Asp Pro Asp Asp Val Glu Gly Cys Asn
234      260      265      270
236 Arg Ala Met Asp Phe Met Leu Gly Trp His Gln His Pro Thr Ala Cys
237      275      280      285
239 Gly Asp Tyr Pro Glu Thr Met Lys Lys Ser Val Gly Asp Arg Leu Pro
240      290      295      300
242 Ser Phe Thr Pro Glu Gln Ser Lys Lys Leu Ile Gly Ser Cys Asp Tyr
243 305      310      315      320
245 Val Gly Ile Asn Tyr Tyr Ser Ser Leu Phe Val Lys Ser Ile Lys His
246      325      330      335
248 Val Asp Pro Thr Gln Pro Thr Trp Arg Thr Asp Gln Gly Val Asp Trp
249      340      345      350
251 Met Lys Thr Asn Ile Asp Gly Lys Gln Ile Ala Lys Gln Gly Gly Ser
252      355      360      365
254 Glu Trp Ser Phe Thr Tyr Pro Thr Gly Leu Arg Asn Ile Leu Lys Tyr
255      370      375      380
257 Val Lys Lys Thr Tyr Gly Asn Pro Pro Ile Leu Ile Thr Glu Asn Gly
258 385      390      395      400
260 Tyr Gly Glu Val Ala Glu Gln Ser Gln Ser Leu Tyr Met Tyr Asn Pro
261      405      410      415
263 Ser Ile Asp Thr Glu Arg Leu Glu Tyr Ile Glu Gly His Ile His Ala
264      420      425      430
266 Ile His Gln Ala Ile His Glu Asp Gly Val Arg Val Glu Gly Tyr Tyr
267      435      440      445

```

## RAW SEQUENCE LISTING

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:14

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

```

269 Val Trp Ser Leu Leu Asp Asn Phe Glu Trp Asn Ser Gly Tyr Gly Val
270      450                      455                      460
272 Arg Tyr Gly Leu Tyr Tyr Ile Asp Tyr Lys Asp Gly Leu Arg Arg Tyr
273 465                      470                      475                      480
275 Pro Lys Met Ser Ala Leu Trp Leu Lys Glu Phe Leu Arg Phe Asp Gln
276                      485                      490                      495
278 Glu Asp Asp Ser Ser Thr Ser Lys Lys Glu Glu Lys Lys Glu Ser Tyr
279                      500                      505                      510
281 Gly Lys Gln Leu Leu His Ser Val Gln Asp Ser Gln Phe Val His Ser
282                      515                      520                      525
284 Ile Lys Asp Ser Gly Ala Leu Pro Ala Val Leu Gly Ser Leu Phe Val
285                      530                      535                      540
287 Val Ser Ala Thr Val Gly Thr Ser Leu Phe Phe Lys Gly Ala Asn Asn
288 545                      550                      555                      560
292 <210> SEQ ID NO: 3
293 <211> LENGTH: 5
294 <212> TYPE: PRT
295 <213> ORGANISM: Artificial Sequence
297 <220> FEATURE:
298 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
299 peptide
301 <400> SEQUENCE: 3
302 Thr Phe Asn Glu Pro
303 1 5
306 <210> SEQ ID NO: 4
307 <211> LENGTH: 5
308 <212> TYPE: PRT
309 <213> ORGANISM: Artificial Sequence
311 <220> FEATURE:
312 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
313 peptide
315 <220> FEATURE:
316 <221> NAME/KEY: MOD_RES
317 <222> LOCATION: (1)
318 <223> OTHER INFORMATION: Ile or Val
320 <400> SEQUENCE: 4
W--> 321 Xaa Thr Glu Asn Gly
322 1 5
325 <210> SEQ ID NO: 5
326 <211> LENGTH: 534
327 <212> TYPE: PRT
328 <213> ORGANISM: Arabidopsis thaliana
330 <400> SEQUENCE: 5
331 Met Tyr Ser Lys Lys Asn Ser Phe Gly Arg Ser Asp Phe Pro Glu Gly
332 1 5 10 15
334 Phe Leu Phe Gly Thr Ala Ser Ser Ala Tyr Gln Tyr Glu Gly Ala Arg
335 20 25 30
337 Asn Glu Ala Pro Arg Gly Glu Ser Val Trp Asp Thr Phe Val Arg Lys
338 35 40 45

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 09/12/2006  
PATENT APPLICATION: US/10/506,670A      TIME: 15:27:15

Input Set : F:\G1309US3.txt  
Output Set: N:\CRF4\09122006\J506670A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; Xaa Pos. 1

VERIFICATION SUMMARY

DATE: 09/12/2006

PATENT APPLICATION: US/10/506,670A

TIME: 15:27:15

Input Set : F:\G1309US3.txt

Output Set: N:\CRF4\09122006\J506670A.raw

L:321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0